S.No. 41(R)

159

OFFICE OF THE CHIEF ELECTORAL OFFICER,
ANDHRA PRADESH
Ground Floor, South 'H' Block, Secretariat, Hyderabad-500022

Letter No.1471/Elecs.B/A1/2018-3 Dated: 25.04.2018

To
Sri Sandeep Saxena,
Deputy Election Commissioner,
Election Commission of India,
Nirvachan Sadan,
Ashoka Road,
New Delhi – 110 001 (w.e)

Sir,

Sub:- Collection of data during the National Electoral Roll Purification and Authentication Programme (NERPAP) – Meeting by the Deputy Election Commissioner on 28.4.2018 at Hyderabad – Methodology of seeding of Aadhaar data, breakup of Aadhaar data collected and present stage of storage – Information furnished - regarding.

Ref:- From the ECI Lr.No.23/1/~~LET~~/ECI/FUNC/ERD-ER/NERPAP/2015 (Vol.VI), dt.20.04.2018

--:o0o:--

With reference to the letter cited, I furnish herewith the following information in respect of Aadhaar seeding during NERPAP, 2015 in Andhra Pradesh & Telangana States as desired therein:

(i) The method of collection, feeding and seeding of Aadhaar data adopted during NERPAP (Annexure-I)

(ii) District-wise breakup of Aadhaar data collected, fed and seeded through organic and inorganic methods (Annexure-II)

The collected Aadhaar data has been maintained in a separate table duly linking with the EPIC ID AC-wise. With UID No.(AC → EPIC-ID → UID)

Yours faithfully,
Sd/-
CHIEF ELECTORAL OFFICER

Copy to the Chief Electoral Officer, Telangana (w.e)

//Forwarded :: By Order//

SECTION OFFICER

(A) 160

Collection, Feeding and seeding of Aadhaar with Electoral Roll Database

Approach:

State Resident Data Hub (SRDH) is a comprehensive portal for managing demographic (name, gender, age, photograph and address) for facilitating "UniqueID" identity authentication. The demographic information in this portal is sourced from the UIDAI. The SRDH Portal and search engine has been used for seeding UID (Aadhaar number) in the Electoral roll database.

Methodology

CEO Office supplied / shared the Electoral Roll / EPIC database with the SRDH application. The application match the electoral roll / EPIC entry with Aadhaar / Ration card / or any other departmental data with wider coverage of population above 18 years of age and seed the Aadhaar number in the Electoral Roll. For the 50% and above matched cases, the Aadhaar number have been incorporated against the elector in the pdf provided to BLO/ERO for verification. For EPIC having less than 50% matching with UID, Aadhaar has not been incorporated against their EPIC number. The list has been generated in pdf format polling booth-wise duly sorting as per house number as being done for the electoral roll. These lists have been made available online in the EPIC-Aadhaar Seeding Portal and the same has been downloaded by the ERO using log-in credentials. The downloaded lists have been printed and supplied to the AEROs / BLOs for field verification.

Procedure:

- In the matched cases (matching 50% and above) as well as in mismatched cases, the BLO gone for door to door verification
- BLO verified the correctness of the Aadhaar number already incorporated in the list.
- Wherever the UID number is not available, the BLO collected the copy of Aadhaar Card from the elector or noted down the Aadhaar number from the original card and mention in the report.

- In case, the elector is not available / shifted his residence, suitable remarks such as dead, shifted, door lock, duplicate have been made by the BLO duly enquiring from the neighbours.

- In case of door lock cases or absence of any adult member at the time of his visit, BLO visited that house at least three times on different occasions during morning and evening hours, and not during day time. If even after three such visits, he finds the house locked or is not able to meet any adult member of the household, then he pasted a sticker on the door, which, inter-alia, should mention:

  "I had come to verify your name in the electoral roll and to collect the Aadhaar number to link with the electoral roll, but you are not available. Please contact me .....(name)..... (Mobile No.) at ..... (Polling Station Address) on _____ between 10.00 a.m. to 5.00 p.m or upload your Aadhaar number by using any of the four following methods

  Seeding through Portal

  SMS

  Call Center

  Mobile app

- Even after three such visits and pasting of sticker, if the elector has turned up, the BLO reports the matter to his Supervisor with a specific remarks that the elector is not the ordinary resident of the address given / shifted / dead / duplicate.
- One Supervisor has been appointed for every 10 to 15 BLOs.
- The Supervisor re-verifed 25% of the electors again and reported to the AERO / ERO.

Action to be taken by the AERO /ERO:

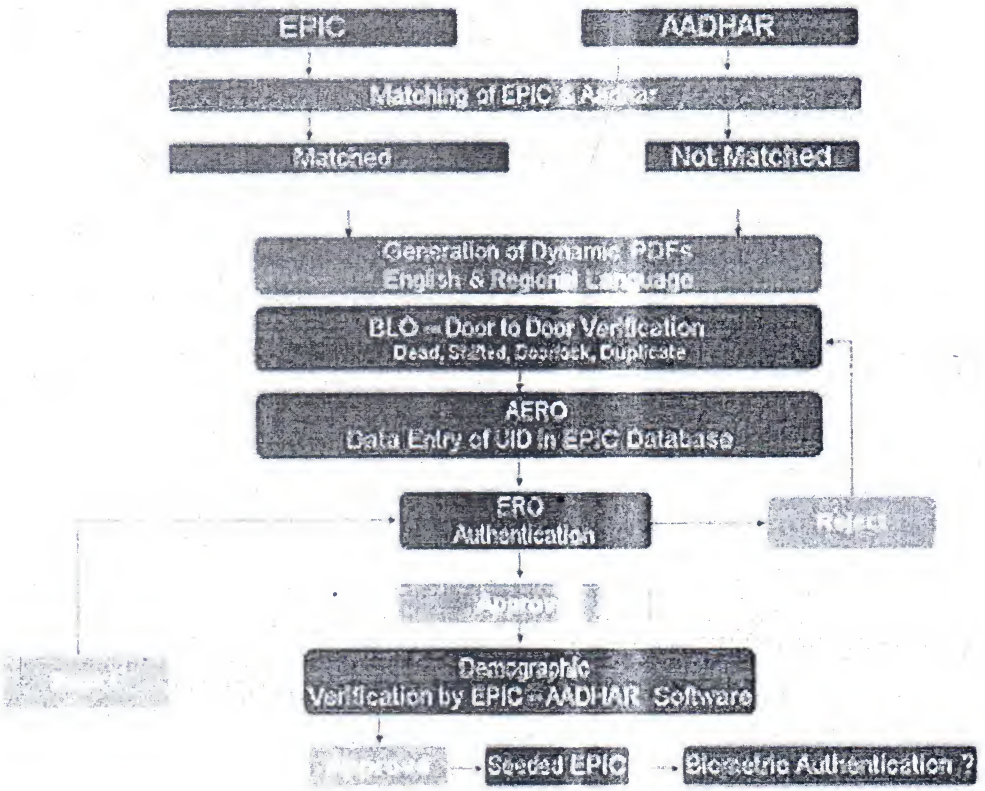
- Computer Operator, Computer and internet connectivity have been provided by the AERO for online data entry.
- Existing computers already provided for Permanent DPLs at each Mandal headquarter have been utilized for the purpose
- AERO taken up supervisory check to the extent of 1% of the entries. The

entries checked by Supervisors and AERO are different from each other duly ensuring that the same household is not visited again and again for verification.

- For the multiple/duplicate entries in the electoral rolls, notices have been issued to the electors under Section 22 of R.P.Act, 1950 and full procedure of verification of facts followed by placing on notice board for a week based on the physical verification reports of the BLOs etc.
- ERO also prepared (under Rule 21A of RER, 1960) a list of the names of dead persons and the persons who may have ceased to be ordinarily residents of the constituency and of persons who are otherwise not entitled to be registered in the roll.
- ERO shall then exhibited on the notice board of his office a copy of the list together with a notice as to the time and place at which the question of deletion of these names from the roll would be considered. After considering any verbal or written objection that was preferred, the Electoral Registration Officers decided whether all or any of the names should be deleted from the roll.
- After verifying following the above mentioned procedure, Aadhaar number has been seeded by ERO in the EPIC database in respect of each elector.
- Once this seeding is completed, demographic verification has automatically run and rejected cases have again gone to the respective Booth and visible in the BLO login and these have been re-verified and submitted for seeding by BLO. If the re-submitted Aadhaar number is the same as seeded earlier then this seeding has not gone to the demographic authentication and have been 'Force seeded' into the EPIC database.
- Once the entire Demographic authentication is completed, SRDH run a de-duplication process over the entire State to find out the duplicates based on the following:
    - Same Aadhaar seeded against multiple EPICs
    - Since both the databases – EPIC and UID are having photos, the duplicates can be easily identified.

- Lists of duplicates have been generated based on the Aadhaar number incorporated in the electoral rolls and said lists have been made available to the EROs in pdf format for further verification.
- Such duplicates have been removed duly following the procedure already mentioned supra.

- The final product will be an electoral roll without multiple entries.



MATCHING

EPIC-AADHAAR matching solution compares the Election department demographic fields Vs UID captured demographic fields and calculates the matching scores in % based on the matched demographic fields to avoid the manual entry of the AADHAAR number against the Electoral as part of the Seeding

Following are the matching solution steps

1. Study the Demographic Data

First step is to study the EPIC department shared demographic fields and generate the data availability and accuracy reports

2. Identify the Matching Fields

Following are the identified demographic fields used for calculating the matching scores for EPIC Data

S.No	UID Data - Fields	EPIC Data - Fields
1	Citizen Name	Electoral Name
2	Care of Name	Father Name/Mother Name/Care of Name
3	Gender	Gender
	Age/Date of Birth	Age/Date of Birth
4	Door Number/Flat Number	Door Number/Flat Number
5	Area/Locality	Area/Village/Locality/ Polling Center name
6	Town/City/Village	Polling Center/Area/Village/Block/Town
7	District	District
8	AADHAAR Photograph	EPIC Photograph

3. **Customize the matching algorithms based on data availability and accuracy**

Following are the high level matching algorithms used and are customized to suit the requirements of the department based on available demographic data and accuracy levels

A. Soundex Algorithm

Soundex is a phonetic algorithm for indexing names by sound, as pronounced in English. The goal is for homophones to be encoded to the same representation so that they can be matched despite minor differences in spelling.

This algorithm is customized to suit the Indian naming conventions and it can handle most of the possible scenarios while identify the matches as well as for calculating the matching scores

Ex: The words "siva" and "shiva" are spelled phonetically same

B. Levenstein distance algorithm

Levenshtein distance is a string metric for measuring the difference between two sequences

Ex: The levenstein distance for the strings "siva" and "shiva" is '1'

C. Short and Long Name algorithm

This algorithm designed for comparing short string and long strings.

Ex: "S Krishna" and "Siva Krishna"

D. String shuffling algorithm

Shuffled strings are same in this case.

Ex: "Siva Krishna "and "Krishna Siva"

4. **Calculate the Matching Scores**

The matching scores would be calculated based on match's accuracy level and weightages given accordingly. The matching scores can be calculated based on following scenarios as per the matching field values.

a. Strings (Names)
    i. Based on the matching characters
b. Numerical Values (Data of Birth)

i. Based on the time period (Month/Year)

c. Area

i. With in/Out of Village/Polling Center

ii. With in/Out of Block

iii. With in/Out of District

5. Aggregate the matches and group them into different buckets

EPIC Matching aggregate the matches into 5 buckets based on the matching scores

1. 100% Matching
2. 99 – 85% matches
3. 85- 70% matches
4. 70 – 50% matches
5. < 50% matches (Name matches (Electoral Vs UID name)

SEEDING

AADHAAR Number Seeding is the process of placing AADHAAR number against Elector in the Electoral roll. It is achieved through online or offline mode. Following are the methodologies adapted to seed UID Numbers into EPIC database.

A. **In-Organic Seeding:** In-organic seeding is a method by which EPIC department demographic fields are compared with the equivalent fields in UID KYR database and generate the matching score based on accuracy level. For One-One matches would be automatically seeded into the EPIC database. One to many matches can be seeded by the departments based field verification feedback.

B. **Organic Seeding:** Organic seeding is a method which allows seeding the Electoral UID Number in to the department's databases in Electoral presence where the BLO/ERO voluntarily reached to the Electoral and seed the AADHAAR based on the AADHAAR letter.

mode with UDIAI Authentication service. The demographic verification based on

i. Name
ii. Gender
iii. Date of Birth

In addition to this the application can also facilitate the seeded beneficiaries with biometric authentication services as part validation and verification services using AADHAAR Authentication services

c. **Fraudulent Data identification (Duplicates/Dead/Migrated etc)**

The application generate the fraudulent electors reports online

d. **Training & Support**

Provides the technical assistance and training to the Election department project stakeholders on need basis

--:o0o:--

AADHAAR SEEDING IN ELECTORAL ROLL DATABASE - STATUS AS ON 24.08.2015

District Name	Electoral Count	Collection of Aadhaar				BLO Verification Status																		ERO Verification Status					
		Mobile	Self	Call Center	SMS	Aadhar Linked	%	Duplicate	%	Ineligible	%	Death	%	Shifted	%	Door Lock	%	Not Enrolled	%	Total BLO Verified Count	%	BLO To be Verified Count	Approved	Rejected	Pending	%	Total Aadhaar Linked	%
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
KRISHNA	2915374	956	21545	242	17276	2630903	90.24	6,022	0.21	1,590	0.05	14,177	0.49	113877	3.91	46,611	1.60	95,301	3.27	2908481	99.76	6893	2892567	9,592	6,322	0.22	2593930	88.97
VIZIANAGARAM	1674356	54	14784	249	9507	1431662	85.51	6,788	0.41	148	0.01	12,010	0.72	65,587	3.92	135845	8.11	20,885	1.25	1672925	99.91	1431	1641537	30,543	845	0.05	1400174	83.62
PRAKASAM	2413284	48	59164	1828	18325	2054547	85.13	8,844	0.37	1,896	0.08	15,186	0.63	109587	4.54	154048	6.38	65,314	2.71	2409422	99.84	3862	2368866	19,305	21,251	0.88	1988470	82.40
EAST GODAVARI	3033793	564	43100	558	22955	3158228	82.38	24,042	0.63	3,813	0.10	48,193	1.26	279400	7.29	236251	6.16	78,070	2.04	3827997	99.85	5796	3787032	37,610	3,355	0.09	3105268	81.00
WEST GODAVARI	2903294	126	39740	3259	20750	2393511	82.44	11,594	0.40	2,005	0.07	27,555	0.95	152620	5.26	221196	7.62	87,403	3.01	2895084	99.74	7410	2847784	34,276	13,824	0.48	2335020	80.43
KURNOOL	2693959	44	54905	5321	15211	2199625	81.65	30,834	1.14	6,213	0.23	15,216	0.56	213456	7.92	193873	7.20	27,391	1.02	2686608	99.73	7351	2561574	107834	17,200	0.64	2118013	78.62
SRIKAKULAM	2009110	317	18892	1379	13446	1594273	79.35	16,786	0.84	1,466	0.07	28,698	1.43	159991	7.96	173828	8.65	29,285	1.46	2004327	99.76	4783	1969315	23,843	11,169	0.56	1566587	77.97
YSR-CUDDUPAH	1977780	93	42326	898	14188	1576883	79.73	14,136	0.71	1,859	0.09	12,442	0.63	111119	5.62	202434	10.24	52,972	2.68	1971845	99.70	5935	1946583	19,640	5,622	0.28	1541494	77.94
GUNTUR	3618212	385	84647	1426	29671	2794899	77.25	29,808	0.82	6,512	0.18	52,130	1.44	284476	7.86	325396	8.99	116272	3.21	3609493	99.76	8719	3492455	51,150	65,888	1.82	2739363	75.71
CHITTOOR	3070142	265	47802	767	17575	2258391	73.56	31,452	1.02	3,375	0.11	35,851	1.17	210544	6.86	429573	13.99	88,790	2.89	3057976	99.60	12166	2993684	62,729	1,563	0.05	2212050	72.05
ANANTAPUR	3022326	377	82751	2874	30882	2057718	68.08	19,797	0.66	3,776	0.12	25,671	0.85	225991	7.48	639693	21.17	41,347	1.37	3013993	99.72	8333	2962395	29,460	22,138	0.73	2004029	66.31
NELLORE	2243744	30	29587	9783	13110	1513362	67.45	28,853	1.29	6,529	0.29	34,654	1.54	264151	11.77	251168	11.19	142225	6.34	2240942	99.88	2802	2157975	61,083	21,884	0.98	1425527	63.53
VISAKHAPATNAM	3434953	80	95194	2686	40347	2243440	65.31	48,211	1.40	3,093	0.09	57,282	1.67	515636	15.01	463187	13.48	56,983	1.66	3387832	98.63	47121	3190854	180793	16,185	0.47	2150854	62.62
ANDHRA PRADESH	35810327	3339	634437	31270	263243	27907442	77.93	277167	0.77	42275	0.12	379065	1.06	2706435	7.56	3473103	9.70	902238	2.52	35687725	99.66	122602	122602	34812621	667858	1.86	27180779	75.90

Number of Electors linked the Aadhaar with EPIC - District-wise as on 11.08.2015 in Andhra Pradesh & Telangana States

Name of the District		Total Number of Electors	Total Aadhaar Linked	Percentage of linking
1		2	3	4
1	KRISHNA	2915374	2593930	**88.97**
2	VIZIANAGARAM	1674356	1400174	**83.62**
3	PRAKASAM	2413284	1988470	**82.40**
4	EAST GODAVARI	3833793	3105268	**81.00**
5	WEST GODAVARI	2903294	2335020	**80.43**
6	KURNOOL	2693959	2118013	**78.62**
7	SRIKAKULAM	2009110	1566587	**77.97**
8	YSR-CUDDUPAH	1977780	1541494	**77.94**
9	GUNTUR	3618212	2739363	**75.71**
10	CHITTOOR	3070142	2212050	**72.05**
11	ANANTAPUR	3022326	2004029	**66.31**
12	NELLORE	2243744	1425527	**63.53**
13	VISAKHAPATNAM	3434953	2150854	**62.62**
ANDHRA PRADESH		**35810327**	**27180779**	**75.90**
1	NIZAMABAD	1404451	1386347	**98.71**
2	ADILABAD	1591685	1518786	**95.42**
3	KARIMNAGAR	2322994	2180460	**93.86**
4	NALGONDA	2225125	2085570	**93.73**
5	MEDAK	1853538	1657694	**89.43**
6	WARANGAL	2383575	1969651	**82.63**
7	MAHABUBNAGAR	2563609	2107705	**82.22**
8	KHAMMAM	1885838	1453340	**77.07**
9	RANGA REDDY	5468324	2327612	**42.57**
10	HYDERABAD	4101035	1614610	**39.37**
TELANGANA		**25800174**	**18301775**	**70.94**
Total of AP & TS		**61610501**	**45482554**	**73.82**